

УДК 004.52

НЕЙРОМЕРЕЖЕВІ МАТЕМАТИЧНІ МОДЕЛІ ЛЮДИНОМАШИННОГО СПІЛКУВАННЯ ТА РОЗПІЗНАВАННЯ

Кирилов С. О.¹, Кирилова Л. О.², Суворова С. Г.³

¹Одеський національний морський університет

²Одеський національний економічний університет

³Пенітенціарна академія України

Анотація: Дослідження присвячене нейромережевій інтерпретації задачі людиномашинного спілкування та розпізнавання за множиною критеріїв, розглянутої як задачі про призначення. Основна мета полягає в зведенні цієї задачі до стандартної форми, де кількість груп критеріїв дорівнює кількості ранжированих запитів. У дослідженні визначено архітектуру нейронної мережі та запропоновано використання мережі бінарних нейронів, що представляє собою матрицю визначеної розмірності. Запропонована модель ранжировання базується на нейронній мережі, яка містить довільні зворотні зв'язки. Це дозволяє передавати збудження назад до нейрона, що сприяє повторному виконанню його функції. Однак у динамічних нейронних мережах виникає нестійкість, що виявляється у випадковій зміні станів нейронів без досягнення стаціонарних станів. Питання стійкості динаміки таких систем залишається відкритим. Розглянута дискретна нейромережа Хопфілда має такі характеристики: один шар елементів, кожен елемент зв'язується з усіма іншими, але не з самим собою; за один етап оновлюється лише один елемент; елементи оновлюються у випадковому порядку, але кожен оновлюється з однаковою частотою; функція виходу бінарна (значення "0" або "1"). Нейромережа Хопфілда є рекурентною: вихід мережі повторно використовується як введення до досягнення стійкого стану. Після запуску нейромережа змінює свій стан, поступово переходячи до усталеного режиму, що дозволяє ідентифікувати план оцінювання процесу людиномашинного спілкування за множиною критеріїв. Для уточнення результатів використовуються процедури випадкового пошуку. Запропонована енергетична функція мінімізується для забезпечення виконання обмежень та розв'язання задачі. Побудована функція досягає мінімуму лише в станах, що відповідають планам призначень. Визначення параметрів мережі здійснюється шляхом зіставлення отриманих функцій з енергетичною функцією в загальному вигляді. Практична реалізація моделі продемонструвала, що нейромережа Хопфілда може бути успішно застосована для ранжировання запитів у системах людиномашинного спілкування та розпізнавання, забезпечуючи високу точність та ефективність вирішення задач ранжировання.

Ключові слова: нейромережева модель, нейрон, людиномашинне спілкування, енергетичні функція, оптимізуючий функціонал.

NEURONET MATHEMATICAL MODELS OF HUMAN-MACHINE COMMUNICATION AND RECOGNITION

S. Kyrylov¹, L. Kyrylova², S. Suvorova³

¹Odesa National Economic University

²Odesa National Maritime University

³Penitentiary Academy Of Ukraine

Abstract: The study is devoted to the neural network interpretation of the task of human-machine communication and recognition by multiple criteria, considered as a task of assignment. The main goal is to reduce this problem to a standard form, where the number of criteria groups is equal to the number of ranked documents. The study defined the architecture of a neural network and proposed the use of a network of binary neurons, which is a matrix of a certain dimension. The proposed ranking model is based on a neural network that contains arbitrary feedback. This allows the excitation



to be transmitted back to the neuron, which contributes to the repeated performance of its function. However, in dynamic neural networks instability occurs, which is manifested in a random change in the states of neurons without reaching stationary states. The question of stability of the dynamics of such systems remains open. The considered discrete Hopfield neural network has the following characteristics: one layer of elements, each element is connected to all others, but not to itself; only one element is updated per stage; elements are updated in random order, but each is updated with the same frequency; the output function is binary (value "0" or "1"). A Hopfield neural network is recurrent: the output of the network is reused as input until a steady state is reached. After starting, the neural network changes its state, gradually moving to a stable mode, which allows identifying a plan for evaluating the process of human-machine communication according to a set of criteria. Random search procedures are used to refine the results. The proposed energy function is minimized to ensure that the constraints are met and the problem is solved. The constructed function reaches a minimum only in the states corresponding to the assignment plans. The definition of the network parameters is carried out by comparing the obtained functions with the energy function in general form. The practical implementation of the model demonstrated that Hopfield's neural network can be successfully applied to document ranking in human-machine communication and recognition systems, providing high accuracy and efficiency in solving ranking problems.

Keywords: neural network model, neuron, human-machine communication, energy function, optimizing functional.

1 ВСТУП

В останні роки спостерігається стрімкий розвиток технологій штучного інтелекту (ШІ), зокрема нейромережових моделей, які значно змінюють способи взаємодії людини з комп'ютерними системами. Однією з ключових областей застосування нейромереж є людиномашинне спілкування та розпізнавання, яке стає все більш важливим у багатьох сферах життя. Нейромережові математичні моделі забезпечують високий рівень точності та ефективності в розпізнаванні мови, зображень, текстів та інших видів інформації, вони дозволяють створювати інтуїтивно зрозумілі інтерфейси для користувачів, підвищують продуктивність праці та покращують якість обслуговування в різних галузях, таких як медицина, освіта, промисловість та розваги. Актуальність дослідження також зумовлена необхідністю покращення взаємодії людини з машинами в умовах зростаючого обсягу даних та складності завдань. Використання нейромережових моделей для створення більш природних і зрозумілих способів спілкування сприяє підвищенню ефективності та зручності використання технологій. Крім того, дослідження нейромережових моделей у контексті людиномашинного спілкування має велике значення для розвитку технологій розпізнавання емоцій та намірів користувачів, що може бути використане для покращення користувацького досвіду та розробки адаптивних систем, які реагують на потреби користувачів у реальному часі. Таким чином, дослідження нейромережових математичних моделей людиномашинного спілкування та розпізнавання є надзвичайно актуальним і важливим для подальшого розвитку сучасних технологій та їх інтеграції в повсякденне життя.

2 АНАЛІЗ ЛІТЕРАТУРНИХ ДАНИХ ТА ПОСТАНОВКА ПРОБЛЕМИ

Людиномашинне спілкування є багатогранною галуззю, що об'єднує знання з комп'ютерних наук, психології, лінгвістики та інших дисциплін. Від ранніх спроб автоматизувати спілкування за допомогою простих алгоритмів до сучасних моделей на основі нейронних мереж, цей напрямок постійно розвивається. Одним із ключових аспектів є розпізнавання мовлення та тексту, що описано в працях [1, 2].

Застосування нейронних мереж у задачах розпізнавання стало революційним кроком, що значно підвищило точність та ефективність цих процесів. Нейромережі, зокрема конволюційні нейронні мережі (CNN) та рекурентні нейронні мережі (RNN), широко використовуються для розпізнавання зображень, мовлення та текстів [3]. Розвиток глибокого навчання дозволив створювати складні моделі, які можуть навчатися на великих масивах даних та здійснювати точне розпізнавання.

Задачі оптимізації є центральними у багатьох дослідженнях, що стосуються нейромереж. Зокрема, методи оптимізації використовуються для навчання нейронних мереж та покращення їх продуктивності. У роботах [4, 5] розглядаються різні підходи до оптимізації, включаючи градієнтний спуск, стохастичний градієнтний спуск та інші методи.

Питання стійкості є важливим аспектом при розробці динамічних нейронних мереж. У дослідженні [6] розглядаються методи забезпечення стійкості нейромереж, включаючи аналіз спектральних властивостей матриць ваг та застосування стабілізуючих функцій активації. Сучасні дослідження акцентують увагу на інтеграції нейронних мереж у системи людиномашинного спілкування та розглядаються приклади використання нейронних мереж для створення чат-ботів, систем розпізнавання емоцій та адаптивних інтерфейсів. У роботах [7, 8] аналізуються різні моделі та алгоритми ранжування, зокрема з використанням нейронних мереж та

особлива увага приділяється питанням оптимізації ранжування та інтеграції різних критеріїв оцінки. Огляд літератури показує, що нейромережеві математичні моделі є потужним інструментом для вирішення задач людиномашинного спілкування та розпізнавання. Сучасні дослідження спрямовані на підвищення точності, стабільності та ефективності цих моделей. Подальші дослідження мають на меті інтеграцію нових методів оптимізації, аналізу стійкості та застосування нейромереж у різних прикладних контекстах.

3 ЦІЛЬ ТА ЗАДАЧІ ДОСЛІДЖЕННЯ

Метою дослідження є розробка та аналіз нейромережевих математичних моделей, які дозволяють вирішувати завдання людиномашинного спілкування та розпізнавання, зокрема, через застосування нейромережі Хопфілда. Дослідження спрямоване на визначення архітектури нейронної мережі, параметрів її функціонування, а також розробку методів для забезпечення стабільності динаміки системи та досягнення точних і ефективних результатів ранжування запитів за безліччю критеріїв.

Задачі дослідження:

- розглянути задачі людиномашинного спілкування та розпізнавання як задачі про призначення, що враховують множину критеріїв;
- побудувати нейромережеву модель ранжування запитів у системах людиномашинного спілкування та розпізнавання.

4 РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

Розглянемо нейромережевому інтерпретацію задачі людиномашинного спілкування та розпізнавання за безліччю критеріїв як задачі про призначення, за умови зведення задачі про призначення до стандартної форми (число груп критеріїв дорівнює числу ранжированих запитів). Визначимо архітектуру нейронної мережі, і введемо в розгляд мережу бінарних нейронів, що представляє собою матрицю розмірністю $n \times n$, де $n = N - M$ – число груп критеріїв. За основу моделі ранжировання може бути взята нейронна мережа (рис. 3.), що містить довільні зворотні зв'язки, за якими передане збудження повертається до даного нейрона, і він повторно виконує свою функцію [9].

У динамічних нейронних мережах нестійкість проявляється в блукає зміні станів нейронів, що не приводить до виникнення стаціонарних станів. У загальному випадку відповідь на питання про стійкість динаміки довільної системи з зворотними зв'язками є вкрай складною і до теперішнього часу є відкритою (рис. 1.).

Нехай використовується дискретна нейромережа Хопфілда має наступні характеристики:

- 1). Один шар елементів (вхідні елементи, що представляють вхідний зразок, не враховуються).
- 2). Кожен елемент зв'язується з усіма іншими елементами, але елемент не зв'язується з самим собою.
- 3). За один етап оновлюється тільки один елемент.
- 4). Елементи оновлюються у випадковому порядку, але в середньому кожен елемент повинен оновлюватися в одній і тій самій мірі (частоті).
- 5). Введення елемента обмежено значеннями “0” або “1”, тобто функція виходу бінарна [10].

Нейромережа Хопфілда є рекурентною в розумінні того, що для кожного вхідного зразка вихід мережі повторно використовується в якості введення до тих пір, поки не буде досягнуто стійкого стану. Відповідним чином організована (запрограмована)

нейромережа після «запуску» змінює свій стан, поступово переходячи в усталений режим. За цим режимом ідентифікується результат – план оцінювання процесу людиномашинного спілкування та розпізнавання безліччю груп критеріїв, який може не збігатися з точним рішенням. Для уточнення результату, зазвичай, використовуються процедури випадкового пошуку [11].

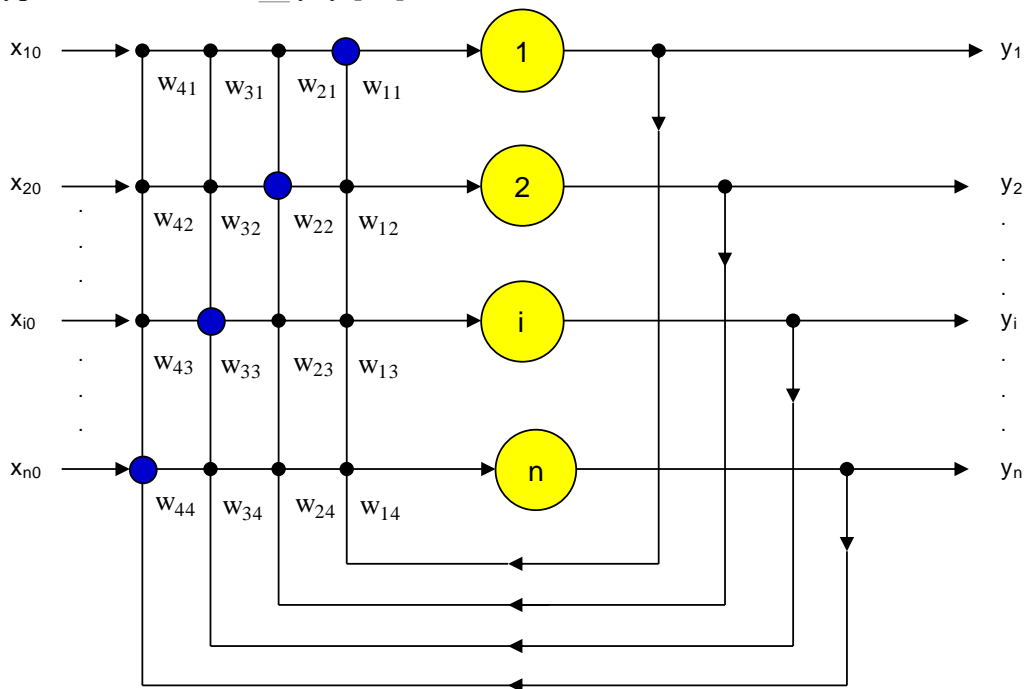


Рис. 1. Нейромережа Хопфілда в задачі ранжування процесу людиномашинного спілкування та розпізнавання (x – вхідні пошукові запити в межах окремої системи; y – виходи нейрона, w – величини зміщень та синаптичних зв'язків)

Зручно вважати, що нейромережа Хопфілда не має вхідних елементів, так як вхідний вектор просто визначає початкові значення активності елементів. Елементи оновлюються тоді, коли всі елементи передадуть свої значення активності за наявними зваженим зв'язків, після чого обчислюється сума доданків (тобто береться скалярний доданок). Значення активності елемента виходить на основі використання деякого правила активізації. Кожній цілочисельній змінній x_{ij} – поставимо у відповідність вихідний сигнал ij -го нейрона u_{ij} , що стоїть в i -му рядку і в j -му стовпці матриці мережі

$$(x_{ij} = 1) \Leftrightarrow (u_{ij} = 1), \forall i, j \in \overline{1, N} \quad (1)$$

Сукупність збуджених нейронів інтерпретується як план призначень. Відповідно до (10), інтерпретуємо обмеження (2) і цільову функцію (3), в результаті отримуємо:

$$\sum_{j=1}^n u_{ji} = 1, \forall i \in \overline{1, n}, \quad (2)$$

$$\sum_{i=1}^n u_{ji} = 1, \forall j \in \overline{1, n}, \quad (3)$$

$$\sum_{j=1}^n \sum_{i=1}^n u_{ji} = n, \quad (4)$$

$$\Phi(u) = \sum_{j=1}^n \sum_{i=1}^n u_{ji} r_{ji} \rightarrow \max, \quad (5)$$

де u_{ji} – значення виходу нейромережі Хопфілда; r_{ji} – значення матриці продуктивності елементи якої r_{ji} представляють собою релевантність сторінки пошуку номером j щодо критерію (групи критеріїв) з номером i .

Сконструємо енергетичну функцію $E^0(u)$, мінімізація якої забезпечує виконання обмежень (2)-(3) і розв'язок задачі (4). Побудуємо її у вигляді

$$E^0(u) = E_{\psi}^0(u) + E_{\Phi}^0(u), \quad (6)$$

де останній доданок забезпечує оптимізацію функції вартості і з точністю до константи $F > 0$ однозначно визначається наступним чином

$$E^0(u) = -F \sum_{j=1}^n \sum_{i=1}^n u_{ji} r_{ji}, \quad (7)$$

а перший доданок забезпечує виконання обмежень і може бути побудований декількома способами. Відповідно до першого з них цей компонент енергетичної функції, що конструюється, має вигляд

$$E_{\psi}^0(u) = \frac{A}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{v \neq i}^n u_{ji} u_{ji} + \frac{B}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu \neq i}^n u_{ji} u_{ji} + \frac{C}{2} \left(\sum_{j=1}^n \sum_{i=1}^n u_{ji} - n \right)^2, \quad (8)$$

де A , B і C – позитивні константи. Перший доданок приймає мінімальне і нульове значення лише в тому випадку, якщо кожен рядок матриці $\{u_{ij}\}$ містить не більше однієї одиниці, другий доданок приймає мінімальне нульове значення, якщо кожен стовпець даної матриці містить не більше однієї одиниці, нарешті, третій доданок приймає мінімальне нульове значення, якщо у всій матриці $\{u_{ij}\}$ міститься рівно n одиниць.

Побудована функція $E_{\psi}^0(u)$ досягає свого мінімуму в усіх станах, які відповідають сукупності обмежень (2)-(3) і представляють собою план призначень. Відповідно до другого способу побудови цього компонента енергетичної функції, що конструюється, будемо мати

$$E_{\psi}^0(u) = \frac{A}{2} \sum_{j=1}^n \left(\sum_{i=1}^n u_{ji} - 1 \right)^2 + \frac{B}{2} \sum_{i=1}^n \left(\sum_{j=1}^n u_{ji} - 1 \right)^2, \quad (9)$$

де перший доданок приймає мінімальне нульове значення тільки в тому випадку якщо в будь-якому рядку матриці $\{u_{ij}\}$ буде тільки один збуджений нейрон, а другий доданок – якщо в будь-якому стовпці цієї матриці буде рівно один збуджений нейрон.

В цілому вказана функція приймає мінімальне нульове значення тільки у станах, які відповідають обмеженням (1)-(2) і представляють собою плани призначень. Підсумовуючи функцію (7) з функцією (8) або (9), сконструємо енергетичну функцію в завершеному вигляді

$$E^0(u) = \frac{A}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{v \neq i}^n u_{ji} u_{jv} + \frac{B}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu \neq i}^n u_{ji} u_{j\mu} + \frac{C}{2} \left(\sum_{j=1}^n \sum_{i=1}^n u_{ji} - n \right)^2 - F \sum_{j=1}^n \sum_{i=1}^n u_{ji} r_{ji}, \quad (10)$$

або

$$E^0(u) = \frac{A}{2} \sum_{j=1}^n \left(\sum_{i=1}^n u_{ji} - 1 \right)^2 + \frac{B}{2} \sum_{i=1}^n \left(\sum_{j=1}^n u_{ji} - 1 \right)^2 - F \sum_{j=1}^n \sum_{i=1}^n u_{ji} r_{ji}. \quad (11)$$

Визначимо параметри мережі, зіставивши одну з отриманих функцій з енергетичною функцією, записаної в загальному вигляді

$$E(u, T, I) = -\frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu=1}^n \sum_{\nu=1}^n T_{ji\mu\nu} u_{ji} u_{\mu\nu} + \sum_{j=1}^n \sum_{i=1}^n u_{ji} I_{ji}, \quad (12)$$

де $T_{ji\mu\nu}$ – коефіцієнт зв'язку між входом ij -го нейрона і виходом $\mu\nu$ -го; I_{ji} – зміщення ij -го нейрона.

У цьому виразі для енергетичної функції мережі навмисне опущений часовий параметр в зв'язку з тим, що при визначенні синапсів і зовнішніх зсувів він не відіграє істотної ролі як для мереж з дискретним часом, так і для мереж з безперервним часом. Більш того, даним виразом слід користуватися при визначенні параметрів синтезованих мереж, як з дискретними, так і з безперервними станами. Підставою для цього служить той факт, що енергетичні функції мереж з дискретними і з безперервними станами відрізняються тільки наявністю у останніх інтегрального доданка, яке як від значень синапсів, так і від зовнішніх зсувів в явному вигляді не залежить [12].

Для того, щоб визначити параметри мережі відповідно до побудованої енергетичної функції (10), наведемо вираз для цієї функції до виду (12)

$$E^0(u) = \frac{A}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\nu \neq i} u_{ji} u_{j\nu} + \frac{B}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu \neq i} u_{ji} u_{\mu i} + \frac{C}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu=1}^n \sum_{\nu=1}^n u_{ji} u_{\mu\nu} - Cn \sum_{j=1}^n \sum_{i=1}^n u_{ji} - F \sum_{j=1}^n \sum_{i=1}^n u_{ji} r_{ji} + \frac{C}{2} n^2. \quad (13)$$

Прирівняємо коефіцієнти при лінійних і квадратичних членах останнього виразу і енергії (12). Останній доданок з розгляду можна виключити, так як він не залежить від стану мережі спілкування. Зіставлення лінійних членів дозволить визначити значення зовнішніх зсувів, а зіставлення квадратичних членів дозволить визначити синаптичні зв'язки між нейронами. Аналіз першого доданка сконструйованої енергетичної функції свідчить про те, що будь-який нейрон мережі повинен мати синаптичні зв'язки з коефіцієнтом $-A$ з усіма нейронами однойменного з ним рядка (умова $\mu=i$) крім самого розглянутого нейрона (умова $\nu \neq j$).

Другий доданок диктує наявність зв'язків з коефіцієнтом $-B$ між нейронами однойменного стовпця (умова $\nu=j$) крім власного зворотного зв'язку (умова $\mu \neq i$). Третя складова свідчить про те, що всі нейрони мережі пов'язані один з одним синапсами з коефіцієнтами $-C$. Скориставшись параметром $-\delta_{ji}$, сформуємо результуючий вираз для синаптичних зв'язків мережі у вигляді

$$T_{ji\mu\nu} = -A\delta_{i\mu}(1-\delta_{iv}) - B\delta_{iv}(1-\delta_{i\mu}) - C = -A\delta_{i\mu} - B\delta_{iv} + (A+B)\delta_{i\mu}\delta_{iv} - C, \quad i, j, \mu, \nu \in \overline{1, n}. \quad (14)$$

Аналіз четвертого і п'ятого доданків сконструйованої енергетичної функції свідчить про те, що на всі нейрони мережі необхідно подавати зовнішні зміщення у вигляді

$$I_{ji} = -Cn - Fc_{ji}, \quad i, j \in \overline{1, n}. \quad (15)$$

Як правило, в практичних завданнях приймають $F=I$ і $A=B$, тоді всі ненульові зв'язки мають однакову вагу, яка дорівнює A . Крім того, аналізуючи вирази (14) і (15), можна помітити, що наявність глобальних зв'язків з коефіцієнтом C кожного нейрона з

кожним в кінцевому стані мережі, відповідає плану призначень, забезпечує подачу на будь-який нейрон з боку всіх інших сумарного сигналу, який дорівнює Cn , який компенсується постійним зміщенням Cn . Отже, для спрощення структури синапсів мережі глобальними зв'язками з вагою C і частиною зсуву Cn в першому наближенні можна знехтувати [13].

Бажана модель нейронної мережі процесу людиномашинного спілкування та розпізнавання містить матрицю з $n \times n$ нейронів, на кожен з яких подається зовнішнє зміщення, яке дорівнює відповідній продуктивності $I_{ji} = -r_{ji}$, а вихідний сигнал будь-якого нейрона u_{ji} , з коефіцієнтом A , подається на входи всіх нейронів однойменних з ним рядки і стовпці.

Інший варіант параметрів мережі для оптимального плану оцінювання процесу людиномашинного спілкування та розпізнавання можна отримати, використовуючи сконструйовану енергетичну функцію у вигляді (14). Аналогічно проведеній вище процедури наведемо цей вираз до нового виду (15)

$$E^0(u) = \frac{A}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{v=1}^n u_{ji} u_{jv} + \frac{B}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu=1}^n u_{ji} u_{\mu i} - A \sum_{j=1}^n \sum_{i=1}^n u_{ji} - B \sum_{j=1}^n \sum_{i=1}^n u_{ji} - F \sum_{j=1}^n \sum_{i=1}^n u_{ji} r_{ji} + \frac{n}{2}(A+B). \quad (16)$$

Порівняємо коефіцієнти при лінійних і квадратичних членах, відкинувши останній доданок. В результаті отримуємо:

$$T_{ji\mu v} = -A\delta_{i\mu} - \delta_{iv}, \\ I_{ji} = -(A+B) - Fr_{ji}, i, j, \mu, v \in \overline{1, n}. \quad (17)$$

Крім розглянутого функціоналу, можна використовувати різні комбінації функцій (6) і (7) для конструювання енергетичної функції $E^0(u)$. В результаті будемо отримувати різні варіанти параметрів нейромережі. При побудові нейроподібні мережі з безперервними станами необхідно забезпечити умови знаходження точок спокою в кутах n -мірного куба її простору станів. Іншим способом забезпечення суворої бінарності вихідних сигналів нейронів в стійких станах є додавання до енергетичної функції, що конструюється, додаткової складової, що досягає мінімального значення на станах мережі, в яких вихідні сигнали нейронів беруть значення "0" або "1" [14]. Прикладом такої функції може служити наступна

$$E_p^0(u) = \frac{G}{2} \sum_{j=1}^n \sum_{i=1}^n u_{ji} (1 - u_{ji}) = \frac{G}{2} \sum_{j=1}^n \sum_{i=1}^n u_{ji} - \frac{G}{2} \sum_{j=1}^n \sum_{i=1}^n u_{ji}^2, \quad (18)$$

де $G > 0$ – константа. Додавши цей доданок до раніше побудованої енергетичної функції, наприклад у вигляді (19), отримуємо

$$E^0(u) = \frac{A}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{v \neq i}^n u_{ji} u_{jv} + \frac{B}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu \neq i}^n u_{ji} u_{\mu i} + \frac{C}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu=1}^n \sum_{v=1}^n u_{ji} u_{\mu v} - \frac{G}{2} \sum_{j=1}^n \sum_{i=1}^n u_{ji}^2 - Cn \left(\sum_{j=1}^n \sum_{i=1}^n u_{ji} - F \sum_{j=1}^n \sum_{i=1}^n u_{ji} r_{ji} + \frac{C}{2} n_i^2 \right). \quad (19)$$

звідки визначимо наступні параметри нейромережі пошуку

$$\begin{aligned}
 T_{ji\mu\nu} &= -A\delta_{i\mu}(1-\delta_{iv}) - B\delta_{iv}(1-\delta_{i\mu}) - C + G\delta_{i\mu}\delta_{iv} = \\
 &= -A\delta_{i\mu} - B\delta_{iv} + (A+B+G)\delta_{i\mu}\delta_{iv} - C, \\
 I_{ji} &= \frac{G}{2} - Cn - Fr_{ji}, j, \mu, \nu \in \overline{1, n}.
 \end{aligned}
 \tag{20}$$

Об'єднавши вирази (20) і (16), можна сконструювати енергетичну функцію у такому вигляді

$$\begin{aligned}
 E^0(u) &= \frac{A}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{v=1}^n u_{ji}u_{jv} + \frac{B}{2} \sum_{j=1}^n \sum_{i=1}^n \sum_{\mu=1}^n u_{ji}u_{\mu i} - \frac{G}{2} \sum_{j=1}^n \sum_{i=1}^n u_{ji}^2 + \frac{G}{2} \sum_{j=1}^n \sum_{i=1}^n u_{ji} - \\
 &- A \sum_{j=1}^n \sum_{i=1}^n u_{ji} - B \sum_{j=1}^n \sum_{i=1}^n u_{ji} - F \sum_{j=1}^n \sum_{i=1}^n u_{ji}r_{ji} + \frac{n}{2}(A+B),
 \end{aligned}
 \tag{21}$$

звідки шукані параметри мережі визначаються наступним чином:

$$\begin{aligned}
 T_{ji\mu\nu} &= -A\delta_{i\mu} - B\delta_{iv} + G\delta_{i\mu}\delta_{iv}, \\
 I_{ji} &= \left(A + B - \frac{G}{2} \right) - Fr_{ji}, j, \mu, \nu \in \overline{1, n}.
 \end{aligned}
 \tag{22}$$

Ранжування запитів людиномашинного спілкування та розпізнавання у разі єдиного критерію значущості, тобто в граничному випадку – $V=I$, реалізується при наступному припущенні: є безліч значень релевантності списку запитів, знайдених в результаті запиту: $\{r_i\}$, $i \in N=1\dots, n$. В результаті ранжирування, числа розставляються в порядку зростання індексів $j = \pi(i)$ так що, $\forall j \in 1\dots(n-1), r_j \leq r_{j+1}$.

Пара чисел $\{r_k, r_l\}$ була обрана довільно. Аналогічні докази можна надати і до будь-якої іншої пари чисел. Зокрема, умова (1) буде справедливою для різних пар сусідніх елементів. В силу того, що будь-яка перестановка може бути представлена суперпозицією транспозиції сусідніх елементів, максимальне значення лінійного функціоналу буде досягтися на перестановці, в якій всі пари сусідніх елементів відсортовані [15]. Таким чином, вихідна задача пошуку сортуючої перестановки зведена до оптимізації лінійного функціоналу на множині всіляких перестановок, розв'язання якої визначається у вигляді (2). Для розв'язання такого завдання за допомогою нейроподібні мережі наведемо її нейромережевому інтерпретацію [16]. Уявімо довільну підстановку $j = \pi(i)$ у вигляді матриці нейронів $[V_{ij}]$ розмірністю $n \times n$ наступним чином

$$(j = \pi(i)) \Leftrightarrow (V_{ij} = 1), \forall i, j \in \overline{1, n}.
 \tag{23}$$

На змістовному рівні збуджений стан нейрона $V_{ij}=1$ відповідає тому факту, що i -й елемент вихідної множини значень релевантності r , займає j -е місце в підстановці $\pi(i)$. У будь-якому рядку і кожному стовпці такої матриці повинен бути рівно один збуджений нейрон, отже справедливі обмеження виду:

$$\begin{aligned}
 \sum_{i=1}^n V_{ij} &= 1, \forall j \in \overline{1, n}, \\
 \sum_{j=1}^n V_{ij} &= 1, \forall i \in \overline{1, n}.
 \end{aligned}
 \tag{24}$$

Нейромережева інтерпретація оптимізуючого функціоналу буде представлена наступним чином

$$\phi(V) = \sum_{j=1}^n a_j \sum_{i=1}^n r_j V_{ij} \rightarrow \max_M \quad (25)$$

Це завершальний етап побудови алгоритму.

5 ОБГОВОРЕННЯ РЕЗУЛЬТАТІВ ДОСЛІДЖЕННЯ

Порівнявши останні вирази з результатами нейромережевої інтерпретації задачі про призначення у вигляді (1)-(4), приходимо до висновку про те, що оптимізаційна постановка задачі людиномашинного спілкування та розпізнавання тотожна постановці задачі про призначення за умови, що елементи матриці продуктивностей останньої визначаються як

$$c_{ij} = a_j r_i \quad i, j \in 1, \dots, n. \quad (26)$$

Отже, подальше побудова нейромережі людиномашинного спілкування та розпізнавання буде повністю збігатися з побудовою мережі для розв'язання задачі про призначення. При цьому параметри мереж для вирішення як першої, так і другої задачі будуть визначатися одними і тими ж виразами. Зокрема, скориставшись параметрами мережі для вирішення задачі про призначення з урахуванням умови (22), отримаємо параметри мережі для вирішення задачі сортування в наступному вигляді:

$$\begin{aligned} T_{j\mu\nu} &= -A\delta_{i\mu} - \delta_{i\nu}, \\ I_{ji} &= -(A+B) - Fr_{ji}, \quad j, \mu, \nu \in \overline{1, n}. \end{aligned} \quad (27)$$

Використовуючи інші вирази для визначення параметрів мережі, вирішує завдання про призначення, можна отримати родину нейромереж для розв'язання задачі сортування даних в системі людиномашинного спілкування та розпізнавання.

6 ВИСНОВКИ

Дослідження показало, що задачі людиномашинного спілкування та розпізнавання можуть бути ефективно інтерпретовані як задачі про призначення. За умови зведення задачі про призначення до стандартної форми, де кількість груп критеріїв дорівнює кількості ранжируваних запитів, вдається досягти оптимального розподілу ресурсів і підвищити точність розпізнавання. Визначення архітектури нейронної мережі, що включає мережу бінарних нейронів, показало, що для досягнення стійких станів необхідно враховувати специфічні характеристики та зв'язки між нейронами. Використання енергетичної функції для мінімізації обмежень і цільової функції дозволяє забезпечити стабільність та оптимізацію функціонування нейромережі, що є критично важливим для динамічних нейронних мереж. Практична реалізація та верифікація запропонованої моделі показала, що нейромережа Хопфілда може бути успішно застосована для ранжування запитів у системах людиномашинного спілкування та розпізнавання. Розроблені методи дозволяють отримувати точні результати, які відповідають сукупності обмежень та забезпечують високий рівень продуктивності при вирішенні завдань ранжування.

7 ЕТИЧНІ ДЕКЛАРАЦІЇ

Автори не мають відповідних фінансових чи нефінансових інтересів, які слід розкривати.

Література

1. Gan G., Ma Ch., Wu J. Data Clustering: Theory, Algorithms and Applications. Philadelphia, Pennsylvania: SIAM, 2007. 455 p.
2. Alsabhan W. Human–Computer Interaction with a Real-Time Speech Emotion Recognition with Ensembling Techniques 1D Convolution Neural Network and Attention. *Sensors*. 2023. №23(3). p. 1386. <https://doi.org/10.3390/s23031386>.
3. McAfee A., Brynjolfsson E. Machine, Platform, Crowd: Harnessing Our Digital Future. New York: W.W. Norton & Company, 2017. 340 p.
4. Македон В. В., Байлова О. О. Планування і організація впровадження цифрових технологій в діяльність промислових підприємств. Науковий вісник Херсонського державного університету. Серія «Економічні науки». 2023. Випуск 47. С. 16-26. DOI: 10.32999/ksu2307-8030/2023-47-3.
5. Yashchenko V.A. Neural-like growing networks in the development of general intelligence. Neural-like element (P. I) Mathematical machines and systems. 2022. N4. P. 15-36.
6. Дзямучич М. І., Шматковська Т. О. Вплив сучасних інформаційних систем і технологій на формування цифрової економіки. Економічний форум. 2022. №2. С. 3-8.
7. da Silva I.N., Spatti D.H., Flauzino R.A., Bartocci Liboni L.H., dos Reis Alves S.F. Artificial Neural Networks: A Practical Course. Springer International Publishing, 2016. 307 p.
8. Македон В. В., Холод О. Г., Ярмоленко Л. І. Модель оцінки конкурентоспроможності високотехнологічних підприємств на засадах формування ключових компетенцій. Академічний огляд. 2023. № 2 (59). С. 75-89. DOI: 10.32342/2074-5354-2023-2-59-5.
9. Hryshko V., Zinchenko I. Synergy of AI and business as a factor of management efficiency in the context of technology development. Science Journal «Economics and Region». 2023. Vol. 4(91). pp. 223-229. doi: [https://doi.org/10.26906/EiR.2023.4\(91\).3216](https://doi.org/10.26906/EiR.2023.4(91).3216).
10. Jin K. H., McCann M. T., Froustey E., Unser M. Deep convolutional neural network for inverse problems in imaging. IEEE Transactions on Image Processing. 2017. 26. pp. 4509–4522.
11. Makedon V., Dzeveluk A., Khaustova Y., Bieliakova O., Nazarenko I. Enterprise multi-level energy efficiency management system development. International Journal of Energy, Environment, and Economics. 2021. Volume 29. Issue 1. pp. 73-91.
12. Li Ruichao, Nawi Abdullah, Kang, Myoung. Human-machine Translation Model Evaluation Based on Artificial Intelligence Translation. EMITTER International Journal of Engineering Technology. 2023. №11. pp. 145-159. 10.24003/emitter.v11i2.812.
13. Zhao Ting. Artificial Intelligence in Mathematical Modeling of Complex Systems. EAI Endorsed Transactions on e-Learning. 2024. №10. 10.4108/eetel.5256.
14. Scibilia A., Pedrocchi N., Fortuna L. Human Control Model Estimation in Physical Human–Machine Interaction: A Survey. *Sensors*. 2022. №22(5). P. 1732. <https://doi.org/10.3390/s22051732>.
15. Pronina O., Piatyko O. The recognition of speech defects using convolutional neural network. CTE Workshop Proceedings [Online]. 2023. №10. pp. 153-166. Available from: <https://doi.org/10.55056/cte.554>.
16. Dawodi M., Baktash J.A., Wada T., Alam N., Joya M.Z. Dari Speech Classification Using Deep Convolutional Neural Network. 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). 2020. pp.1-4. Available from: <https://doi.org/10.1109/IEMTRONICS51293.2020.9216370>.

References

1. Gan, G., Ma, Ch., Wu, J. (2007). *Data Clustering: Theory, Algorithms and Applications*. Philadelphia, Pennsylvania: SIAM.
2. Alsbhan, W. (2023). Human–Computer Interaction with a Real-Time Speech Emotion Recognition with Ensembling Techniques 1D Convolution Neural Network and Attention. *Sensors*, 23(3), 1386. <https://doi.org/10.3390/s23031386>.
3. McAfee, A., & Brynjolfsson, E. (2017). *Machine, Platform, Crowd: Harnessing Our Digital Future*. New York: W.W. Norton & Company.
4. Makedon, V. V., Bailova, O. O. (2023). Planuvannya i orhanizatsiya vprovadzhennya tsyfrovyykh tekhnolohiy v diyal'nist' promyslovykh pidpryyemstv [Planning and organizing the implementation of digital technologies in the activities of industrial enterprises]. *Naukovyy visnyk Khersons'koho derzhavnoho universytetu. Seriya «Ekonomichni nauky»*, Issue 47, 16-26. DOI: 10.32999/ksu2307-8030/2023-47-3. [in Ukrainian].
5. Yashchenko, V.A. (2022). Neural-like growing networks in the development of general intelligence. *Neural-like element (P. I) Mathematical machines and systems*, N4, 15–36.
6. Dzyamulych, M. I., Shmatkovs'ka, T. O. (2022). Vplyv suchasnykh informatsiynykh system i tekhnolohiy na formuvannya tsyfrovoyi ekonomiky [Influence of modern information systems and technologies on the formation of the digital economy]. *Ekonomichnyy forum*, №2, 3-8. [in Ukrainian].
7. da Silva, I.N., Spatti, D.H., Flauzino, R.A., Bartocci Liboni, L.H., dos Reis Alves, S.F. (2016). *Artificial Neural Networks: A Practical Course*. Springer International Publishing.
8. Makedon, V. V., Kholod, O. H., Yarmolenko, L. I. (2023). Model' otsinky konkurentospromozhnosti vysokotekhnolohichnykh pidpryyemstv na zasadakh formuvannya klyuchovykh kompetentsiy [The model of assessing the competitiveness of high-tech enterprises based on the formation of key competencies]. *Akademichnyy ohlyad*, 2(59), 75-89. DOI: 10.32342/2074-5354-2023-2-59-5 [in Ukrainian].
9. Hryshko, V., Zinchenko, I. (2023). Synergy of ai and business as a factor of management efficiency in the context of technology development. *Science Journal «Economics and Region»*, 4(91), 223-229. doi: [https://doi.org/10.26906/EiR.2023.4\(91\).3216](https://doi.org/10.26906/EiR.2023.4(91).3216).
10. Jin, K. H., McCann, M. T., Froustey, E., Unser, M. (2017). Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26, 4509–4522.
11. Makedon, V., Dzeveluk, A., Khaustova, Y., Bieliakova, O., Nazarenko, I. (2021). Enterprise multi-level energy efficiency management system development. *International Journal of Energy, Environment, and Economics*, Volume 29, Issue 1, 73-91.
12. Li, Ruichao & Nawi, Abdullah & Kang, Myoung. (2023). Human-machine Translation Model Evaluation Based on Artificial Intelligence Translation. *EMITTER International Journal of Engineering Technology*, 11, 145-159. 10.24003/emitter.v11i2.812.
13. Zhao, Ting. (2024). Artificial Intelligence in Mathematical Modeling of Complex Systems. *EAI Endorsed Transactions on e-Learning*, 10. 10.4108/eetel.5256.
14. Scibilia, A., Pedrocchi, N., Fortuna, L. (2022). Human Control Model Estimation in Physical Human–Machine Interaction: A Survey. *Sensors*, 22(5), 1732. <https://doi.org/10.3390/s22051732>.
15. Pronina, O. and Piatyko, O. (2023). The recognition of speech defects using convolutional neural network. *CTE Workshop Proceedings [Online]*, 10, 153–166. Available from: <https://doi.org/10.55056/cte.554>.
16. Dawodi, M., Baktash, J.A., Wada, T., Alam, N. and Joya, M.Z. (2020). Dari Speech Classification Using Deep Convolutional Neural Network. 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), 1-4. Available from: <https://doi.org/10.1109/IEMTRONICS51293.2020.9216370>.

Кирилов Сергій Олександрович

Одеський національний морський університет,
к.ф.-м.н, доцент
вул. Мечникова, 34, м. Одеса, Україна, 65029
kyrylovserhii@gmail.com



Кирилова Людмила Олександрівна
Одеський національний економічний університет,
к.ф.-м.н, старший викладач
вул. Преображенська, 8, м. Одеса, Україна, 65000
kirilovaludmilaalex@gmail.com
ORCID: 0009-0009-4577-242X

Суворова Світлана Геннадіївна
Пенітенціарна академія України,
к.е.н, доцент,
вул. Гонча, 34, м. Чернігів, Україна, 14000
svetlanasuvorovach@gmail.com
ORCID: 0000-0001-7089-2524

Для посилань:

Кирилов С. О., Кирилова Л. О., Суворова С. Г. Нейромережеві математичні моделі людиномашинного спілкування та розпізнавання. Механіка та математичні методи, 2024.Т. VI. № 2. с. 147–159

For references:

S. Kyrylov, L. Kyrylova, S. Suvorova. (2024). Neuronet mathematical models of human-machine communication and recognition. Mechanics and mathematical methods. VI (2). 147–159.